	Typ	Hits	Search Text	DBs	Time Stamp
16	BRS	52	first adj (document$1 file$1) same second adj (document$1 file$1) same (temporary intermediate) with (stor$4 record$3)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	2004/12/27 11:17
17	BRS	21	first adj (document$1 file$1) same second adj (document$1 file$1) same (temporary intermediate) with (stor$4 record$3) and compar$8	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	2004/12/27 10:32
18	BRS	2	"6658626".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	2004/12/27 10:32
19	BRS	11	S85 not S83	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	2004/12/27 11:11
20	BRS	19	first adj (document$1 file$1) same second adj (document$1 file$1) same (temporary intermediate) with (stor$4 record$3) and @rlad<="20020111"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	2004/12/27 11:14

	Type	Hits	S arch Text	DBs	Time Stamp
21	BRS	147	first adj (document$1 file$1) same second adj (document$1 file$1) same (temporary intermediate)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	2004/12/27 11:17
22	BRS	47	first adj (document$1 file$1) same second adj (document$1 file$1) same (temporary intermediate) and @rlad<="20020111"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	2004/12/27 15:24
23	BRS	47	(document$1 file$1) with compar$8 same (temporary intermediate) and "715"/$.ccls.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	2004/12/27 16:45

Detailed Description Text - DETX (10):

At a step 114 the program stores data relating to any
matches found in the index file after queried for a current
hash value. In the described embodiment, a list of <hash
value, position> pairs are stored in a **temporary** file. The
index file stores positions of each hash value. Thus, at
step 114 if a current hash value is found in the index
file, the position and value stored in the index file
matching the current hash value is stored in the **temporary**
file until all the substrings in the query document (as
described in step 108) have been hashed and searched for in
the index file. In the described embodiment, a position
value encodes (i.e. it does not explicitly state) the name
of the document and an offset within that file where the
hash value begins. Thus, this position value performs as an
absolute position or address within the collection or
corpus of documents insofar that it can be used to go
directly to a position within a document regardless of
where that document resides in the corpus of documents. If
the collection of **documents being compared** against each
other are expected to be dissimilar, step 114 will normally
result in small amounts of matching data or none at all.
However, this depends on the nature of the collection of
documents being compared.

Detailed Description Text - DETX (37):

A similar procedure is applied to the P positions in the
sorted list. First, the list is sorted based on p, instead
of 0. The program then checks for overlaps in P by using
position data in the B-tree. Similarly, in other preferred
embodiments, overlaps in P can be determined by **comparing**

the difference between $p.sub.i - p.sub.i+1$ to 1 since the position pairs have been segmented into <u>documents</u> and the program is checking for overlaps within a single <u>document</u>. For those overlaps that have the same offset as overlaps in the 0 positions, the information stored in the <u>temporary</u> data structure is used to replace the overlapping P passages. Since other position pairs may contain the P value being changed, the P value in those pairs are changed as well to keep the value consistent. In the described embodiment, the <u>temporary</u> data structure maps hash values of segments to positions of those segments in a document.

Current US Original Classification - CCOR (1):

<u>715/526</u>

Current US Cross Reference Classification - CCXR (1):

<u>715/511</u>

Current US Cross Reference Classification - CCXR (2):

<u>715/530</u>